



USAID
FROM THE AMERICAN PEOPLE



PERFORMANCE & IMPACT EVALUATION (P&IE)

Semi-Annual Report, October 2014

31 OCTOBER 2014

This publication was produced at the request of the United States Agency for International Development. It was prepared independently by NORC at the University of Chicago.

PERFORMANCE & IMPACT EVALUATION (P&IE) SEMI-ANNUAL REPORT, APRIL 2014

31 October 2014

PN 7384; USAID Contract N0: AID-617-C-12-00006

PRESENTED TO:

USAID/Uganda

Joseph Mwangi

PRESENTED BY:

NORC at the University of Chicago

Jeffrey Telgarsky

Senior Vice President, International Projects

4350 East-West Highway, 8th Floor

Bethesda, MD 20814

Telephone: (301) 634-9413

Fax: (301) 634-9301

DISCLAIMER

The author's views expressed in this publication do not necessarily reflect the views of the United States Agency for International Development or the United States Government.

CONTENTS

HIGHLIGHTS	3
DESCRIPTION OF ACTIVITIES	4
A. Impact Evaluation Activities.....	4
Data Quality Assessment.....	4
Advice to SHRP Result 2 Team on options for dealing with a deviation from an IRB protocol.....	5
B. Performance Evaluation and CLA Activities	7
B.1 Systematic observation and documentation of project implementation activities as inputs for the performance evaluations.	7
B.2 Collaboration, Learning and Adaptation (CLA) Activities	8
B.3 Midterm Performance Evaluation	10
C. Risks to the Impact Evaluation.....	10
ANNEX 1: excerpt of memo to SHRP TEAM on general observations of ENUMERATOR TRAINING RESULT 1 CLUSTER 1 ROUND 3 AND CLUSTER 2 ROUND 2 EGRA DATA COLLECTION	12
ANNEX 2: FEEDBACK PRESENTED TO THE SHRP TEAM FOLLOWING OBSERVATION OF ENUMERATOR TRAINING AND PILOT TEST FOR RESULT 1 CLUSTER 1 ROUND 3 AND CLUSTER 2 ROUND 2 EGRA DATA COLLECTION.....	16
A. Overview	16
A. Appreciative feedback.....	16
C. Issues and Recommendations	17
Assessor Training.....	19
Field implementation of EGRA.....	20
ANNEX 4: CHALLENGES TO THE IMPACT EVALUATION, AS PRESENTED IN SEMI-ANNUAL REPORTS PRESENTED IN JUNE 2013, OCTOBER 2013, and april 2014	39

NORC at the University of Chicago (NORC), in partnership with the Panagora Group, is pleased to submit to USAID/Uganda this Semi-Annual Report recording progress made on the P&IE project between May 1 – October 31, 2014.

HIGHLIGHTS

During this reporting period, the NORC/Panagora team:

- Completed a data quality review of the EGRA Cluster 2 Round 1 dataset. This data, which will be used, along with Cluster 2 Round 2 data, to measure the impact of SHRP's Reading (Result 1) interventions for Cluster 2 students, was collected by the IP in February 2014. NORC received the dataset in May 2014, and reviewed completeness and internal consistency of the dataset in preparation for the impact analysis.
- Completed a data quality review of the KAP 2 dataset. This data along with the KAPI data constitute the baseline data for the impact evaluation of SHRP's HIV/AIDS (Result 2) interventions. The dataset was reviewed for completeness and internal consistency.
- Conducted data quality assessment (DQA) tasks related to the Cluster 1 Round 3 and Cluster 2 Round 2 EGRA for which data collection took place in October 2014:
 - ▶ Reviewed and provided feedback to the Implementing Partner (IP) on EGRA instruments and training manuals, both from the perspective of NORC's role as evaluator and data quality reviewer.
 - ▶ Participated in all parts of enumerator training: the new Data Quality Assessor training on September 22, 2014; the new assessor training on September 24-26; and the main assessor training prior to field work on September 29-October 3.
 - ▶ Travelled to three schools each in Manafwa District and in Wakiso District to observe EGRA data collection activities.
 - ▶ Identified and brought to the attention of the IP and USAID several quality issues with EGRA training and data collection.
- Worked with the SHRP Result 2 Team and provided advice on options for dealing with a deviation from an IRB protocol.
- Completed the Midterm Performance Evaluation. The P&IE Performance Evaluation team, led by our subcontractor, the Panagora group, travelled to Uganda for three weeks in June/July 2014 to conduct interviews and focus group discussions for the Mid-Term Performance Evaluation, and subsequently prepared a detailed "Mid-Term Performance Evaluation Report" for SHRP. The report was submitted to USAID for review and comment on September 30, 2014.
- Continued the ongoing process of third party monitoring and performance feedback with meetings and activities, based on a monthly calendar and an events and assignments tracker to ensure comprehensive monitoring.
 - ▶ Continued to observe RTI/SHRP implementation activities, including leadership and management trainings, school support supervision exercises, and EGRA Assessors training. During this period, eleven events were observed.

- ▶ Continued to record, using our observation tools, information and observations on each activity for the mid-term and final SHRP performance evaluation; and noted appreciative and constructive comments to provide as monthly performance feedback to RTI within the context of the CLA process.
 - ▶ Continued to conduct monthly performance feedback meetings with SHRP Chief of Party and M&E Director, as well as key Results 1 and 2 team members based on observations from SHRP events and activities. During the reporting period we held three feedback meetings (July, August, and September), which are valued and appreciated by RTI/SHRP.
 - ▶ Implemented regular check-in meetings between our in-country staff and the SHRP R1 and R2 leads to allow for increased efficiency and consistency in planned activities monitored on a weekly basis.
- Met with P&IE COR, Joseph Mwangi, in NORC's Bethesda offices on September 23, 2014 to discuss progress on project activities and discuss timing of a workshop to disseminate the Midterm PE and First Annual Impact Evaluation findings to a broad group of stakeholders.

DESCRIPTION OF ACTIVITIES

A. IMPACT EVALUATION ACTIVITIES

Data Quality Assessment

During this reporting period, P&IE staff engaged in various data quality assessment (DQA) tasks related to the Result 1 Cluster 2 Round 2 dataset and to the Result 1 Cluster 1 Round 3 and Cluster 2 Round 2 EGRA data collection, which took place in October 2014. They included the following activities:

- Conducted a data quality review of the Cluster 2, Round 1 EGRA data collected in February 2014. NORC staff noted a few consistency issues particularly in the School Observation data (number of registered students and number of students attending the class conflicting with each other), but overall, we found the data to be of high quality. However, NORC found that the data from three schools which were documented as having been surveyed by the field teams do not appear in the dataset. NORC will follow up with the SHRP team about this issue.
- Conducted a comparison of the sample of schools between Cluster 1 Round 1, Cluster 1 Round 2 and Cluster 1 Round 3 data collections, as well as between Cluster 2 Round 1 and Cluster 2 Round 2 data collections. The lack of proper documentation and comprehensive field reports from the SHRP EGRA data collection make it challenging to ensure that sample selections are respected by the field teams, especially considering the changes to the sampling plans between each round of data collection and the need to replace certain schools during the field period itself. For instance, some schools that had been included in the sample for the Cluster 1 Round 1 data collection could not be re-surveyed for the Round 2 data collection and were therefore replaced during the Round 2 field period. It appeared that for Round 3, the SHRP team decided to re-visit the Round 2 replacements rather than the original Round 1 schools (as would have been the preference of NORC). This decision was made without consulting the NORC evaluation team. As such, NORC conducted a comprehensive comparison between the samples of each round of data collection and will share the results of this exercise with the SHRP team in the month of November.

- Reviewed and provided feedback on data collection instruments (EGRA and learner context instruments, teacher/head teacher survey, classroom observation tool, school inventory) and training manuals for the Result 1 Cluster 1 Round 3 and Cluster 2 Round 2 EGRA data collection and provided written feedback to the SHRP team in September 2014 prior to the enumerators training. In providing feedback on instruments we were cognizant of maintaining consistency in instruments across clusters. For instance, we note that for the vocabulary and word segmenting subtasks, both "no response" and "all correct" can be selected as response options. We had raised this issue previously with RTI who did not seem able to program the question differently to allow for mutual exclusivity of these two different options, as a result we did not raise the issue again. We also noted that the training manuals were significantly improved from previous rounds as roles of DQA and supervisors/enumerators as well as school and sampling procedures are better defined. The SHRP team took into consideration the comments made by NORC on the February 2014 Field Manual for this round of data collection.
- Participated in all parts of enumerator training: the new Data Quality Assessor training on September 22, 2014; the new assessor training on September 24-26; and the all-inclusive assessor training prior to field work, which took place during the week of September 29. NORC's Resident Evaluation Manager participated in the new DQA and assessor trainings, while NORC's Literacy and EGRA Expert, Gaelle Simon, from School-to-School International, traveled to Uganda on September 27-October 12 to be present at the main assessor training. Both traveled to the field to observe data collection during the first few days of the field period. Annex 1 includes excerpts of a memo containing overall feedback on the three training sessions provided to RTI by the NORC/Panagora team. The memo also contains (*in italics*) RTI's responses to observations.

Annex 2 includes a compilation of detailed technical feedback on specific aspects of the training provided to RTI by NORC during the training, either verbally or through email. Of particular concern were several issues related to the size of the training making it difficult to ensure quality control of the assimilation of materials by assessors, erroneous or contradictory information communicated to assessors regarding the administration of EGRA, and translation issues of instructions into local languages.

- Conducted field observations in three schools each (total of six schools) in the Manafwa district (Lumasaaba language) on October 6-8 and in Wakiso District (Luganda language) on October 14-15. These field observations were undertaken by NORC's Resident Evaluation Manager (in both Manafwa and Wakiso) and visiting Literacy and EGRA Expert (Manafwa). Observations resulting from these field visits are also included as part of Annex 2.

Advice to SHRP Result 2 Team on options for dealing with a deviation from an IRB protocol

During this reporting period, the SHRP Result 2 team contacted NORC to clarify several points of confusion regarding the KAP1 and KAP2 data collections in preparation for renewal of their IRB Protocol.

Background description of KAP1 and KAP2: The KAP1 data collection, fielded in June 2013, included Cluster 1 primary and post-primary schools with the exclusion of some boarding schools which could not be included in the sample as students could not obtain signed consent forms from their guardians/parents. As a result, preparations were made to include Cluster 1 boarding schools in the KAP2 data collection, fielded in September 2013, by distributing consent forms with sufficient lead time

prior to data collection such that boarding school students could bring them home during their school break. The KAP2 data collection also included Cluster 2 primary schools but not Cluster 2 post-primary schools as these schools were no longer part of the Result 2 intervention. Furthermore, while the Cluster 1 data collection included treatment and control schools within treatment districts as well as comparison schools in comparison districts, Cluster 2 data collection included treatment and control schools within treatment districts only. The timing of the KAP2 data collection was also brought forward in order to avoid school examination periods which would have jeopardized the ability to reach the target sample size. Both KAP1 and KAP2 data are considered baseline data.

Result 2 Team Questions and NORC Answers: As a result of staff turnover within the Result 2 team as well as of the changes in data collection plans due to the consent form issue in boarding schools and timing due to school examination periods, confusion arose about (1) the rigor of the evaluation and (2) what was proposed in the original IRB submission versus what happened in reality.

Regarding the rigor of the evaluation, NORC clarified that the evaluation design is sound for both Cluster 1 and Cluster 2 schools. We explained that the lack of comparison districts for Cluster 2 schools does not invalidate the research design. The lack of comparison districts only means that the impact of the district-level intervention cannot be estimated. However the impact of the school-level intervention can still be estimated by comparing treatment and control schools within treatment districts. Furthermore, the fact that Cluster 1 boarding schools were assessed at a later point in time than other Cluster 1 schools does not pose problem as the SHRP Result 2 team had assured NORC that the intervention had not yet started (or was in very early stages) at the time of the KAP2 data collection.

Regarding possible mismatches between the original IRB protocol and the actual data collections, it appears that the deviation mainly concerns the number of schools and learners for KAP2, as well as the timing of the KAP2 data collection. The Result 2 team communicated to NORC that the Ugandan IRB, NARC, confirmed that these changes should have been reported prior to the KAP2 data collection but that the Result 2 team had not done so. In late July 2014, NORC recommended that the Result 2 team continue conversations with NARC and submit documentation of deviations from the original protocol as soon as possible in order to resolve all confusion prior to the submission of the IRB protocol renewal for 2015. NORC also recommended that the Result 2 team inform USAID. NORC and the Result 2 team had a follow-up conversation in September 2014 to answer additional questions about the number of post-primary schools included in the sample as well as about the evaluation design.

Current Status: NORC has not received an update from the Result 2 team since the last conversation in September 2014. NORC will be following up with the Result 2 team in order to ensure that all IRB concerns are resolved and that IRB protocols are renewed well in advance of the next round of data collection scheduled to happen in mid-2015.

B. PERFORMANCE EVALUATION AND CLA ACTIVITIES

B.I Systematic observation and documentation of project implementation activities as inputs for the performance evaluations.

The NORC/Panagora team, continued to implement our processes for systematically monitoring and documenting SHRP implementation activities, and our system for carrying out the CLA elements of our contract and providing RTI with performance feedback.

Monthly monitoring routine. During this period, we continued to follow a sequence of information collection, reporting, review, and feedback that is captured in a monthly calendar (see attached sample calendar for details) that includes a monthly in-person P&IE country staff (Resident Evaluation Manager and Sr. HIV/AIDS Specialist) meeting to review the prior month's work and determine content of the monthly report and the performance feedback memo on feedback to RTI based on information from the observation reports, and conduct forward planning; a deadline for P&IE country staff to send a monthly report and draft RTI performance feedback memo to NORC and Panagora; a full P&IE team meeting with NORC, Panagora, and in-country staff; a monthly performance feedback session with RTI; and a Performance Evaluation/CLA coordination meeting (Panagora, in-country staff).

SHRP Events and Assignment Tracker. We continued to maintain and update a SHRP Events and Assignment Tracker which serves as a management tool to ensure coverage of events by our Resident Evaluation Specialist and Senior HIV/AIDS Evaluator and to track submission of all observation reports. The tracker lists the name of the event, date, a description and type of event, duration, observation tool used, the observer assigned to the event, date report submitted by our in-country staff, and if applicable whether an SHRP training report was obtained from RTI. SHRP Events and Assignment Tracker continued to be updated and utilized for forward planning during our monthly P&IE team meetings.

Observation and monitoring tools. During this period our Resident Evaluation Specialist and Senior HIV/AIDS Evaluator continued to successfully use the following our observation and monitoring tools.

P&IE document binder. To facilitate access and review of implementation progress, in particular in anticipation of the imminent mid-term performance evaluations, all reports generated continue to be systematically organized and numbered as soft copies and hard copies into a binder.

Monthly report. During this period we continued to use the updated and more efficient monthly report template and successfully generated detailed and thorough monthly reports.

During this reporting period, our in-country staff attended and observed the following meetings, events, and trainings, and prepared a report on each of them using the appropriate observation and monitoring tool. Each of the reports provided appreciative and constructive observations that were collated and shared with RTI as performance feedback.

- School Support Supervision for Result 2 in low performing primary schools in the Wakiso district, Kakiri CC, Wampewo CC, and Kyaliwajjala CC on July 16th and 17th, 2014
- RTI Uganda Supplier Day (vendors meeting) in Kampala on July 31st, 2014
- Implementing Partners Planning Meeting for Year Three SHRP in Kampala on July 23rd and 24th, 2014
- Master Trainers Leadership and Management Training in Wakiso district and Shimon Core PTC on August 12th, 13th, and 14th, 2014

- TOT Leadership and Management Training of the Bushenyi district trainees in Mbarara district at Bishop Stuart Primary Teachers College on August 20th and 21st, 2014
- TOT Leadership and Management Training of Ngora district trainees in Ngora district at St. Aloysious Ngora Primary Teachers College on August 20th and 21st, 2014
- Phase One of the Head Teachers and Teachers Leadership and Management Training of the Wakiso district trainees in Wakiso district at Shimon Core Primary Teachers College
- EGRA new DQA Training and Field Practice at SHRP offices on Yusuf Lule Road on September 22nd, 2014
- EGRA New Assessors Training and Field Practice at SHRP offices on Yusuf Lule Road on September 24th – 26th, 2014
- EGRA Main Assessors Training in Ndeeba at Pope Paul Memorial Training Center on September 29th – October 3rd, 2014
- HIV Data Collection in program schools to submit to MEEP in Namayumba, Wakiso district, at St. Mathias Primary School on September 24th, 2014

B.2 Collaboration, Learning and Adaptation (CLA) Activities

To implement the CLA component of the P&IE contract, we continued to provide performance feedback to RTI on a monthly basis, with both appreciative and constructive feedback, focusing on elements of performance where real-time feedback will help to strengthen performance and lead to optimal outcomes. The performance feedback continues to be drawn directly from the reports of meetings, events, and activities observed by P&IE in-country staff.

Three months of performance were covered during the reporting period, July, August, and September. In April, there weren't observable events; in May, our Resident Evaluation Manager was on maternity leave; and throughout June, she and our Sr. HIV/AIDS Evaluation Specialist were engaged in the mid-term performance evaluation. RTI continues to express its appreciation for the value of the feedback memos and the subsequent discussions which has given them access to insights that allow them to improve their performance in real time. We are pleased that this exercise continues to be regarded by RTI as a valued and welcome opportunity to improve performance and results achievement as a part of regular implementation. As part of the mid-term performance evaluation, we collated all performance feedback to date into one matrix in which RTI summarized actions taken as a result of the feedback. This is appended herein as Annex 3 and provides evidence substantiating the value of real-time performance feedback and its contribution to adaptive management.

We provide the following illustrative example of appreciative and constructive performance feedback provided during the reporting period, including RTI's response (in italics). Our resident team observed Result 2 Support Supervision in low performing schools in Wakiso District:

Appreciative Feedback. Supervisors were conversant with SHRP Result 2 strategies and approaches. They responded adequately to teachers' questions which included challenges encountered when implementing SHRP recommended activities. Some of the challenges expressed were about the utilization of SHRP tools/registers, the lack of an adequate number of teachers to lead School Family Initiatives (SFIs), restricted school programs which do not allow for SFI activities to be run, the diversity of ages in SFIs which hinders the proper dissemination of some topics in the PIASCY manual, and questions about how to integrate HIV/Health education in classroom lessons

and co-curricular activities. The discussions generated were helpful to schools that were eager to improve on their performance.

During the teachers' discussions supervisors shared best practices and relevant lessons generated by better performing program schools. Some of the best practices shared/lessons learnt were about use of peer leaders in managing some SFIs and schools integrating HIV education in co-curricular activities like sports and debates. Teachers appreciated the knowledge shared and demonstrated interest in applying these best practices.

Field Assistants were attentive in their responses to teachers' questions regarding the program strategies and approaches. This demonstrated good knowledge about the program. It is also evident that they can also answer questions during their regular visits to schools.

Constructive Feedback. There was a discrepancy in the composition and number of supervisors allocated to a school. One school received a team of four officials (two program staff and two Field Assistants). The second school got a team of two officials (one program staff and one Field Assistant). The third school was visited by two officials who were Field Assistants. We observed that the schools visited by teams with program staff benefited more from the visit because the program staff was able to answer any questions raised by the school staff. Unfortunately, the school visited by a team of Field Assistants was left with unanswered questions. Field Assistants were unable to respond to some questions on program administrative issues. For example, teachers wanted to know why they were not paid for carrying out SHRP activities and why SHRP teacher trainings were organized during holidays.

RTI response. We will try to find out what the most common questions teachers are asking and ensure that all staff and Field assistants are able to give the same response. They also had clarifying questions on HIV/AIDS topics. There should be consistency in the composition of support supervision and a concerted effort made to ensure that at least one program staff member is included in every team.

Since these were follow up visits prompted by the negative results obtained from previous support supervisions, we expected the supervisors to refer to previous reports to guide the discussions. However, we observed that schools were taken through new assessments identifying implementation gaps which were already known. This consumed too much of the time needed to support the school and also encroached on teachers' time to attend to class lessons. The Result 2 team should think of better ways of utilizing data collected in previous support supervisions to inform the next visits. This will make the time teams spend in a school more efficient and effective.

RTI response. This was our first time to carry out a formal support supervision process with a check list and report format that we now feed into our M&E system. In the past it was mainly reports written by FA's and staff as part of the travel reports. We will continue however to improve on our data collection tool, feedback this time round was that it was too long.

Teachers also had difficulties using the SFI and Guidance and Counseling (G&C) registers provided by the program. The narrative sections for SFI register were never filled in. These sections required teachers to record lessons learnt and challenges faced during implementation of recommended SHRP HIV activities, and also to compile termly summaries about activities carried out. In addition, the G&C register was never used to record learners counseled on related issues. The fact that teachers have not adequately made records in these registers greatly affects the quality of data collected from schools (i.e., numbers of learners reached by HIV program and ability to pick lessons to improve on future school based HIV interventions).

RTI's response. What we have found is that we constantly have to remind them to fill the registers correctly. This is why we carry out termly support supervision.

Another feedback memo, focusing on assessor training, is presented in Annex 1.

B.3 Midterm Performance Evaluation

During this reporting period, our subcontractor leading performance evaluation work under P&IE, Panagora Group, conducted a Mid-Term Performance Evaluation and subsequently prepared a detailed “Mid-Term Performance Evaluation Report” for SHRP. The report was submitted to USAID on September 30.

The performance evaluation drew from the comprehensive monitoring and observations that the NORC team has been conducting since the inception of the projects (described in Section B.2); extensive reviews of SHRP reports, work plans, PMPs, and other documents; and an intensive round of key informant interviews, focus group discussions, and classroom observations conducted during a two week trip to Uganda by the PE team.

The SHRP Mid-Term Evaluation was organized around six evaluation questions covering design, implementation, results, and sustainability, effective use of funds, and management and learning; a summary of the main findings, conclusions, and recommendations for each question are presented in the Midterm Evaluation report. It was a formative, cross-sectional, descriptive, and analytical performance evaluation employing qualitative methods of data collection, which included document review; key informant interviews (KIs) at the national, district, and school level; focus group discussions; and school and classroom observation. We also drew on the continuous monitoring data and performance feedback from P&IE activities. We gathered quantitative data through review of documents and inclusion of close-ended questions in the tools used for key informants. We employed purposeful sampling of schools and districts. Data quality and analysis was validated through triangulation of multiple sources and stakeholders.

This performance evaluation complements P&IE's impact evaluation; while the impact evaluation measures program impact and the degree to which the end results – in particular, reading outcomes - are being achieved, the performance evaluation provided rich qualitative data on stakeholder perceptions, beliefs, and thinking, which has been analyzed by a cross-disciplinary team of health, education, and development experts to develop findings, conclusions, and recommendations.

The performance evaluation team comprised of Team Leader/Evaluation Specialist, Betsy Bassan; Literacy/Education Evaluation Specialist, Brenda Sinclair; and Senior HIV/AIDS Evaluator, Stella Neema. In addition, the P&IE COP, Alicia Menendez, and NORC's home office P&IE director, Varuni Dayaratna, provided ongoing oversight and support. Country-based support was provided by NORC's Resident Evaluation Manager, Evelyn Namubiru, who also participated in district and school site visits to expand the number of schools visited.

C. RISKS TO THE IMPACT EVALUATION

Below we present an assessment of risks/challenges to the impact evaluation that emerged during this reporting period. Challenges identified prior to the current reporting period, which were presented in previous semi-annual reports, are listed in Annex 6.

- I. Result 1: Possible contamination in the control group. In October, we learned from RTI that Mango Tree Project was working in Otuke, a control district for the SHRP evaluation, and providing their literacy intervention to two control schools in the SHRP sample. This occurred despite careful coordination between Mango Tree and SHRP. The schools in question were replaced for others, however the replacement schools have no baseline and their usefulness is limited. In addition, it is possible that some contamination has already occurred as we do not have any type of control over Mango Tree activities in the district.
- Result 1: Sample size adjustments between rounds. The IP has changed sample sizes of each cohort between rounds. For the first cohort of students (Cluster 1), the February 2013 baseline included 280 schools to allow for analysis of 3 treatment arms, controls in treatment districts, and controls in comparison districts. However, the Cluster 1 Round 2 data collection conducted in October 2013, RTI collected data only in a subsample of treatment schools (168 of the 280), since a decision was made to only focus on one (and not 3) treatment. For Cluster 1, Round 3 in October 2014, however, RTI reverted back to data collection from 280 schools to account for that fact that the 3 treatment arms were implemented in the second year. This use of unbalanced panels does not preclude us from conducting a rigorous evaluation; however, it makes the process less transparent and prevents us from having measurements year by year without loss of information and precision. A similar change happened for Cluster 2. In this case, RTI requested NORC to calculate a sample size large enough to be able to analyze results at the district level. At baseline, in February 2014, data was collected from enough number of schools to calculate impact at district level; however, in for the first follow up in October 2014, the IP decided that district level analysis was not of interest and reduced the data collection to a subsample of the original schools. Although we will not be able to say anything about impact at district level, if properly implemented, this change should not prevent us from analyzing results at language level. In general, NORC recommends following the original samples over time to produce a more streamlined and transparent process and more comparable results across year.

ANNEX I: EXCERPT OF MEMO TO SHRP TEAM ON GENERAL OBSERVATIONS OF ENUMERATOR TRAINING RESULT I CLUSTER 1 ROUND 3 AND CLUSTER 2 ROUND 2 EGRA DATA COLLECTION

MEMO TO: RTI/School Health and Reading Program (SHRP)
FROM: NORC/Panagora Performance and Impact Evaluation (P&IE) Team
DATE: 02 October 2014
SUBJECT: Performance feedback on observed SHRP activities in September 2014

We are pleased to provide our monthly feedback to RTI from observed activities during the month of September 2014. We observed the following activities this month:

1. EGRA new DQAs training and field practice, observed on 22nd September 2014

- *Purpose:* A two-day training and field practice organized for newly appointed EGRA Data Quality Assurance officers (DQAs). The training focused on orienting participants on roles and checklist forms used by DQAs during EGRA field data collection.
- *Content/approach:* Content covered during the training included the following topics: the importance of EGRA, the roles of DQAs during field data collection, team building skills, EGRA classroom lesson observation tool, do and don'ts when observing classroom lessons, the Assessor Observation checklist form, tips on conducting daily debriefs/feedback to teams, and tips on uploading tools to the cloud. The training utilized lectures guided by power point presentations, paired practices and plenary discussions. The training also had a half-day field practice on using the classroom lesson observation tool. The field practice was conducted in Busega Community Primary school located in Kampala district (non-program district).
- *Relationship to work plan:* EGRA DQAs training contributes towards IR 1.5, Programs and policies informed by data and research.

2. EGRA New Assessors training and field practice, observed on 24-26th September 2014

- *Purpose:* It was 3 days training that was concluded with a field practice. The training was meant to orient newly recruited Assessors on EGRA subtasks to prepare them for the main Assessors training (refresher training for old Assessors) which was going to happen the following week.
- *Content/approach:* The training covered the following topics: the importance of EGRA, a review of EGRA subtasks, differences between local language and English EGRA versions, learner context section, and field practice. The training methods included lectures, simulations of EGRA subtasks by trainers, practice in pairs, plenary discussions, and group drillings on EGRA subtasks. The field practice was carried out in Busega Primary school in Kampala district. All languages practiced were from the same school.
- *Relationship to work plan:* EGRA new Assessors training contributes to IR 1.5, Programs and policies informed by data and research.

3. EGRA Main Assessors training, observed on 29th Sept-3rd October 2014

- *Purpose:* A five-day training for newly trained and old Assessors. It was a refresher training workshop preparing both the new and old Assessors for the follow up EGRA field data collection.
- *Content/approach:* The training covered the following topics: the importance of EGRA, Early Grade Reading baseline findings, a review of the five EGRA subtasks, a practice on administration of paper and electronic versions of EGRA, and a discussion of teacher, head teacher, and school facilities survey tools. The training conducted lectures on different EGRA subtasks, allowed participants to practice in pairs, carried out group drillings about EGRA administrations, simulations of EGRA subtasks guided by master trainers and conducted Inter Rater Reliability (IRR) for English and local language subtasks.
- *Relationship to work plan:* EGRA new Assessors training contributes to IR 1.5, Programs and policies informed by data and research.

APPRECIATIVE FEEDBACK

1. EGRA new DQAs training and field practice, observed on 22nd September 2014

- The training was mostly delivered by three trainers who had executed DQA roles during many past EGRA field data collection. They were knowledgeable on DQAs roles and responsibilities. Along with lectures delivered during the training, trainers shared relevant experiences, e.g. on team building, data quality assurance and safety of tablets while in the field. Participants were able to learn from people with ground experiences.
- The training sessions were equally distributed among the trainers; each trainer was allocated specific sessions to lead.
- The trainers practiced team teaching and they supplemented each other's sessions. They were always available in the training room and listened to all sessions.
- Participants had a field practice on using classroom lesson observation tool; they were offered an opportunity to practice with the tool before actual field data collection. Classroom lesson observations are activities expected to be conducted by DQAs only.

2. EGRA New Assessors training (camp) and field practice, observed on 24th - 26th September 2014

- The training was delivered by trainers who had facilitated past EGRAs. They had good background knowledge in conducting EGRA Assessors trainings and were conversant with EGRA content. They delivered logical power point presentations and responded adequately to participants questions or issues of clarification during plenary.
- The training methods were adequate. The training utilized lectures guided by projected power point presentations, plenary discussions happening every after a session, group drillings and paired practices done to allow participants practice with the knowledge/skills attained at end of each session. These techniques encouraged individual and peer learning to take place. More so, the techniques encouraged active participation for all participants during sessions.
- Trainers and SHRP M&E team closely supervised practice sessions and offered required coaching to participants who had not yet mastered the training skills. Training allocated enough time to support individual participants. We discovered many new Assessors were doing well in the main Assessors training that followed.

- Team teaching was emphasized in the training and the major trainers, together with the SHRP M&E team, supplemented the training sessions.
- The training sessions were equally shared among the trainers; each trainer was allocated sessions to lead.
- Attendance of newly trained DQAs who had performed Assessors roles in previous EGRAs was useful. During sessions they shared good experiences on how they implemented Assessors roles in previous EGRAs. They were also strategically placed to support trainers in supervision of practice sessions, and participated in coaching of participants who had not yet mastered the skills.

3. EGRA Main Assessors training, observed on 29th Sept-3rd October 2014

- The four-day training (day 5 was reserved for supervisors only) was well attended. Only 13 assessors were new which allowed for a faster training pace.
- The sessions were well organized and sequenced. The day started with a presentation of the schedule of activities. The Trainers were clearly well experienced in EGRA processes, training for and implementation of EGRA.
- The RTI support team was present and available to help with logistics, analysis of IRR results and changes to Tangerine.
- The language group sessions were lively and engaging. The assessors and the trainers seem to have a good relationship. The new DQAs were using their experience in previous EGRA assessment to lead the teams efficiently. Team members reviewed translations, discussed and corrected ambiguities, did pair practice and provided feedback to each other.
- UNEB observers attended daily. They observed the language group practices, provided feedback and engaged with the assessors.

CONSTRUCTIVE FEEDBACK

1. EGRA new DQAs training and field practice, observed on 22nd September 2014

Although the training managed to accomplish activities listed on the agenda, the first day was packed with many activities which cause some sessions to be rushed to keep up to the agenda schedules.

The training reviewed two checklist forms (lesson observation tool and Assessors observation form). The training allocated 30 minutes for discussions on each tool, which only allowed trainers to read out questions in the forms. This did not allow for effort to be focused on helping participants to acquire knowledge on the administration of the forms, and no other training opportunities were planned to support the participants in mastering the content in the forms. For example, it requires a researcher to have a good background of SHRP reading methodology and Thematic Curriculum in order to carry out classroom lesson observations.

SHRP: One of the hallmarks of our classroom observation tool is that it looks for relatively easily observable, objective teacher actions. It was purposely designed this way so that it could be used by individuals who are not well versed in SHRP methodology. This includes things like “beating a word” (clapping or tapping the syllables of a word) or writing a letter in the air. After the classroom discussion, the new and old DQAs went to observe a classroom using the form. When they returned, they compared notes as to what they had observed and hadn’t observed. The tool also guides the DQA to check lesson plans, learner assessment records and learner exercise books.

2. EGRA New Assessors training and field practice, observed on 24th - 26th September 2014

During the field practice, all Assessors went to one school located in Kampala district, which was a Luganda language speaking region. Although it was agreed to attempt the subtasks in English, many of the learners who were selected were not conversant in English, therefore the non-Luganda speakers had difficulty communicating with such learners. In this case the practice mostly benefited only Assessors who were conversant with Luganda.

SHRP: This was of course a trade off as we could not go to all language areas during the course of the new assessor training. We believed it served its purpose for the assessors to have real learner experience. We are happy to report our new assessors are some of the best in the field now.

3. EGRA Main Assessors training, observed on 29th Sept-3rd October 2014 (Gaelle add any constructive feedback for this activity)

- With such a large group of assessors (over 100), it was difficult to engage participants during plenary sessions and for the review of the English tools.

SHRP: It was a big group but with 1 DQA for every language group (12-16 participants) the venue lent itself to appropriate plenary discussions. It was a large group but we don't see a better alternative. We don't really see any other way in order to ensure consistency of messages. When things were broken down by language group, there was more time for smaller discussions.

- Though very experienced and knowledgeable about EGRA, there was confusion amongst the trainers regarding some aspects of the implementation. It would be beneficial for all trainers to review and clarify all training sessions as a team prior to the training workshop. When posed with a question to which an immediate answer was not available, trainers conferred with each other for confirmation of the answer. This is very important to avoid confusion amongst the assessors.

SHRP: It is impossible to anticipate every question and scenario that will come up in training. As the assessors themselves become more experienced, the questions/scenarios become more complicated. All training sessions were reviewed – in fact, the new assessor training just the week before was a great “dry run” for the larger assessor training program.

- A prior thorough review of the English instructions and derived translations is critical to ensure consistency across the language groups.

SHRP: If you look at notes from every training, this same comment comes up. These instructions are reviewed multiple times. With changing and varied orthographies and speakers, issues inevitably arise. Especially now that the assessors themselves are more experienced with the tools.

ANNEX 2: FEEDBACK PRESENTED TO THE SHRP TEAM FOLLOWING OBSERVATION OF ENUMERATOR TRAINING FOR RESULT 1 CLUSTER 1 ROUND 3 AND CLUSTER 2 ROUND 2 EGRA DATA COLLECTION

A. OVERVIEW

From September 29 to October 2, the NORC/STS (School-to-School International) team observed the EGRA assessor training facilitated by SHRP in Kampala, Uganda. Assessors from Cohort 1 (C1) and Cohort 2 (C2) were gathered at a central training venue for 4 days of training. This assessor training followed that of the Data Quality Assurance team (DQA) and new assessors conducted the week before.

The NORC/STS team also observed the 1-day supervisor training on October 3 and traveled to Mbale on October 5 to conduct observations of the implementation of EGRA. The team worked with the Lumasaba language group in Mbale from October 6 to 8 and in Gomba with the Luganda language group on October 14 and 15. In Mbale, each assessor team was observed at least twice. In Gomba, 3 assessor teams were observed. The school observation schedule was as follows:

Date	Language	District	CC	School Name	EMISNO	
06.oct.2014	Lumasaba	Manafwa	Bubulo	BUBWAYA P.S.	6188	C2
	Lumasaba	Manafwa	Busumbu	BUTIRU DEMO P.S.	6167	C2
07.oct.2014	Lumasaba	Manafwa	Sikusi	BUWAKORO P.S.	680030	C2
	Lumasaba	Manafwa	Sikusi	BUTTA P.S.	6220	C2
08.oct.2014	Lumasaba	Manafwa	Bubulo	BUWAGOGO P.S.	6190	C2
14.oct.2014	Luganda	Wakiso	Nsangi	ST. JUDE NAKASOZI P.S.	0305276232	C1B
15.oct.2014	Luganda	Wakiso	Lake Victoria	CHADWICK NAMATE P.S.	0305274442	T2
	Luganda	Wakiso	Lake Victoria	LAKE VICTORIA	0305274422	T2

Overall, the training of the assessors and the implementation of EGRA was well done. However, the NORC/STS team noted some issues and concerns about the implementation of EGRA which should be addressed prior to the next round of activities for Cohort 3 scheduled for February 2015.

A. APPRECIATIVE FEEDBACK

Of the approximately 110 assessors present for training, only 13 were newly trained on EGRA. All others had participated in prior assessments with C1 or C2. This provided several important advantages

to SHRP including a reduced training period and the opportunity to build on assessors' prior experience and knowledge.

SHRP also elected to work with 8 experienced DQAs. While 4 new DQAs were recruited, they were chosen from the pool of experienced assessors so that all DQAs were highly experienced with the implementation of EGRA thus significantly reducing the learning curve.

The DQAs and assessors seemed well at ease with the use of Tangerine. The ability to make rapid improvements to the test forms on Tangerine is a clear advantage over the paper-based tests as is the ability to transmit and compile data in real-time or on a daily basis. It is important to note that, throughout the training, when translation errors were noticed on Tangerine, the SHRP team was very quick to make necessary changes. The use of Tangerine also facilitated the compilation and review of assessor marking and scoring during the inter-rater reliability exercises and provided an opportunity to the SHRP M&E Specialist to receive data daily from the field, review team progress and provide feedback to DQAs to be addressed during daily debrief sessions with the assessors. The SHRP team makes very good use of the advantages offered by Tangerine during the training and implementation of EGRA.

The daily debrief sessions in the field were very well organized. As a result of their observations in the field and the feedback provided by the M&E Specialist, the DQAs had clear objectives/points to address and organized the sessions accordingly. Assessors seemed very receptive to the feedback provided.

C. ISSUES AND RECOMMENDATIONS

The table below provides a summary of the key issues and recommendations that have emerged from our observations of the assessor training and the field implementation of EGRA. A more detailed explanation is provided in the next section for each of the issues included in the table.

Issues	Recommendations
1. Size of training group too large to manage	a. Organize participants in smaller training groups with no more than 3 language groups combined
2. No piloting included during assessor training	b. Include piloting with children even when training experienced assessors
3. Erroneous or contradictory information communicated to assessors during training	c. At end of each training day, organize daily review session with trainers and DQAs to form consensus on issues raised during the day d. Start each day with a review of the previous day's questions/concerns with assessors
4. Too much lag-time between marking activity and IRR results review	e. Encourage assessors and DQAs to review errors immediately after the marking of each subtask and prior to continuing with the next one
5. Maintaining student focus during assessment	f. Encourage assessors to do a short physical activity with students when moving from one language assessment to the next g. Assessors should carefully choose seating area to avoid distractions

Issues	Recommendations
6. Introduction and consent completed too quickly; script not followed	h. Dedicate more time for practice of the introduction and consent during the training i. Revise script and break it up with prompts to retain student focus
7. Students hesitate to move to next letter/word and wait to be prompted by assessor	j. Clarify instructions so that students continue to read without prompting from assessor
8. Assessors rush through subtask instructions and activities	k. Review either the school level sampling needs or the number of assessors per team
9. Inconsistencies between instructions in local language and English instructions	l. Local language and English instructions should be reviewed side by side m. Local language tools should be used during the training of DQAs and supervisors to flag translation errors prior to the assessor training
10. In subtask 1, the placement of letters in the grid can lead to confusion	n. Review placement of letters after randomization to ensure that a letter is not repeated near itself
11. In subtask 2, the inclusion of “You know that each letter has a sound” in the instructions may lead to confusion	o. Review instructions to eliminate any possible confusion with letter sounds p. Review translations of newly corrected instructions in all local languages to ensure that they are consistent
12. In subtask 2, examples used for student practice are not well done (also relevant for subtasks 1 and 3)	q. Particular attention should to be paid to assessors’ modeling skills with the 1 st examples during training
13. In subtask 2, instructions for the examples do not require assessors to say the word to be segmented twice	r. Instructions should be reviewed so that words in examples are said twice by assessors
14. In subtask 2, local language examples provided include words with 3 or more syllables	s. The word selection for the examples should be reviewed to include only words with 2 syllables t. Longer words, 3 or more syllables, can be used as items 6 to 10 (although this may lead to issues with comparison between rounds)
15. In subtask 3, difficult or long words in the first line	u. The choice of words in the first line should be reviewed to include more simple ones
16. In subtask 5 (English vocabulary), students are confused with the instruction to identify words “around us”	v. Provide more precise instructions
17. In subtask 5 (English vocabulary), students will not point to a stranger’s shoes.	w. Accept as correct when students look under the table at the direction of the assessor’s shoes but do not point
18. In subtask 5 (English vocabulary), the use of the singular “shoe” may be confusing	x. Use plural form of “shoes”
19. In subtask 5 (English vocabulary), placement of pencil to the right of the paper	y. If pencil is placed on the right side of the paper but on the paper (not on the table), the answer should be considered correct
20. In subtask 5 (English vocabulary), the assessor hands the pencil to the student prior to each instruction	z. The pencil should remain with the student

Issues	Recommendations
21. In the questionnaire, some questions do not have “no response” as an answer choice but students remain unresponsive	aa. Add “no response” as an answer choice to all questions
22. In the questionnaire, students don’t seem to understand the term “library”	bb. Include a short description of “library” in the question

Assessor Training

Issue: Though sessions were well designed and well-paced, the number of participants during the plenary sessions proved difficult to manage. With over 110 assessors being trained in the same room, it becomes quasi-impossible for trainers to accurately judge individual assessor’s understanding of rules and processes for implementing EGRA.

Recommendation: **Training participants should be organized in smaller groups.** Multiple training rooms should be secured, within the same location if possible, for smaller group training workshops. It is advised that SHRP group no more than 3 language teams (roughly 36 assessors) to form smaller training groups.

Issue: The majority of assessors was experienced with the implementation of EGRA. However, CI assessors had not implemented EGRA for roughly 1 year prior to this training. Skills become rusty. Focusing the shortened training period on EGRA rules is necessary but so is the need to practice with school children prior to the start of the data collection.

Recommendation: SHRP should consider adding **1 day in the field for teams to practice with school children** in non-sampled schools prior to the start of data collection. This additional pilot day will provide a great opportunity for assessors to brush up on required skills for rapport building and consent in addition to practicing the subtask instructions and rules. It also provides an opportunity for the DQAs to observe each assessor with children and to correct any lingering mistakes or misunderstanding prior to data collection.

Issue: There was some lingering confusion amongst the trainers which resulted in erroneous or contradictory information being communicated to assessors. Though the SHRP M&E Specialist or other trainers tried to address all misinformation, given the size of the training group, not all assessors received the corrected information properly. For example, during field visits, we observed assessors “bringing back” students to letters and words while the students were reading. Assessors would ask a student to return to the 3rd letter, for example, even if the student had continued reading and was well beyond that letter. At issue was the worry that students who read fast and incorrectly would receive an “early stop” too early. SHRP has identified “red flags” in the datasets of which one is the “early stop” given to a student too quickly, i.e. within a few seconds of the start of the activity. These red flags are meant to signal when assessors are not giving students enough time to complete the activity; assessors may be moving students along too fast. During training, assessors were told by one trainer that they should bring students back and give them 3 seconds to read each word/letter. And though the M&E Specialist corrected this, i.e. students should be allowed to continue to read at their own pace, the assessors were not aware of the correction. As a consequence, assessors brought students back to early letters/words and counted 3 seconds before guiding them to the next letter/word. Students however continued to read per the instructions which meant that even though a student was reading letter #9, for example, the assessor would bring that student back to read letter #2 again. This negatively impacted students who lost confidence in their reading ability.

Recommendation: **A review session should be organized at the end of every day with all trainers and DQAs.** Confusions and misunderstandings should be addressed and a common

understanding developed. Prior to the start of each day's training, a short session should be organized with assessors to address any confusion and provide the correct answer. **A Frequently Asked Questions (FAQ) should be developed with questions that arise during trainings and reviewed prior to each new training for future rounds of data collection.** This FAQ should be part of the assessor/supervisor/DQA manual.

Issue: The inter-rater reliability activity conducted by SHRP during training provides great information. However, the lag time between when assessors have marked and the review of the IRR results is too long to provide assessors critical understanding of the reasons that lead them to make marking mistakes.

Recommendation: **During the IRR, assessors should review areas of discrepancies in the marking prior to moving on to the next subtask.** This allows the DQAs and assessors to discuss and understand the nature of the mistakes that are being made and to correct them. It is important to note that IRR is a very effective tool to help choose the best assessors for an activity. In this particular instance, all assessors participating in the training would also be fielded so IRR could not be used to weed out weaker assessors. However, IRR coupled with immediate feedback and discussions about errors remains a very important and powerful training tool.

Field implementation of EGRA

Prior to Administration

Issue: Maintaining student focus throughout the assessment has proven challenging for some assessors. Students were easily distracted, they often looked away from the assessor by the second language assessment (second half of the test), or when they realized that they could not perform the tasks.

Recommendation: During training, it was suggested that assessors do a **quick physical activity with students when transitioning from one language assessment to the next.** No assessor was observed doing this during the field implementation but it is recommended that trainers emphasize the need for a quick mental break during the assessor training. Additional recommendations include: a) **organizing the testing space in private areas**, away from the main school yard, b) adequately spacing the teams from one another so that students do not listen or look at other students and assessors during the test (this also ensures that students do not repeat what they hear from others), and c) using the student's name regularly to maintain their attention.

General Administration

Issue: Assessors rushed through the introduction and consent. The script was not always followed and the activity is not always properly explained to students.

Recommendation: We suggest that **more time be dedicated to practicing the introduction and consent during the training.** The sessions that focused on the introduction and consent were too short and infrequent. Additionally, the script for the introduction and consent is quite long. SHRP may want to **revise this script and break it up with prompts** such as "Do you understand?" or "Okay?" for the assessor to maintain the student's focus but also to have it read more as a conversation than a lecture.

Issue: Students seem to hesitate to move from one letter/word to the next and often waited for assessors to prompt them to move on to the next letter/word. This is particularly relevant for subtasks 1 and 3. When students hesitate, they lose time and don't read as many letters or words as they can within 1 minute.

Recommendation: **Instructions should be clarified.** The inclusion of "I will tell you to continue to the next one" in the latter part of the instructions may be the source of the misunderstanding. It should also be clarified and documented which subtasks have a 3 second rule and which have a 10 second rule.

Issue: Assessors took shortcuts with the subtask instructions (they did not provide the complete instructions) and did not consistently apply the 3 seconds or 10 seconds rules. Assessors seemed to be rushing through the assessments in order to complete the 40 students (30 P1 and 10 P2) that are required per school. The school day is too short to complete the required number of assessments so students in P1 and P2 are often asked to stay after school hours for the targets to be met.

Recommendation: The pressure to meet targets will continue to compel assessors to rush through the process if the number of assessors per team is not increased. It is advisable to **revise the number of assessors per team**. Further, if students are requested to remain after school hours to complete the testing, it is advisable that SHRP provide snacks to the students.

Issue: There were discrepancies in the instructions, in some languages, between the paper version and Tangerine as well as between the languages, i.e. instructions in Lugbarati may differ from instructions in Acoli.

Recommendation: A thorough review of the instructions in English is recommended as well as a review of the translations. **Local language and English instructions should be reviewed side by side** to ensure consistency across all languages. Though we understand this is a lengthy task considering the number of languages used by SHRP, it is critical to achieve uniformity of implementation across the languages. We also recommend that **local language tools be used during the training of DQAs and supervisors** so that translation errors are flagged prior to the assessor training.

Subtask 1: Letter Sounds

			a	
A	r	d	o	
a	b	A		

Issue: The placement of the letters in the grid at times leads to mistakes when the same letter is found in clusters. Both students and assessors have gotten confused about the letter being read since the same letter is repeated nearby. For example, with the letter arrangement to the left, when the student skipped line 2 which started with the letter “A” and went to the first letter “a” in line 3, it took a while for the assessor to realize that line 2 had been entirely skipped. Up until that realization, the assessor was marking “r”, “d”, etc., as incorrect. By the time the

assessor realized he was not on the same line as the student, the assessor had already missed a few letters read by the student (whether correctly or incorrectly).

			a	
A	r	d	o	
a	b	A		

Recommendation: Though the letters for subtask 1 are shuffled and thus randomly placed with each version of the EGRA tool, i.e. from baseline to follow-up, etc., it is important that SHRP ensure that the same letter, whether in upper or lower case, is not repeated near itself in the rows directly under, above, to the left, to the right or caddy-corner. For example, another letter “d” should not be placed in the red cells in the grid to the left. Once the randomization is completed, it is

recommended that **SHRP manually rearrange letters** so that the same letter is not found near itself.

Subtask 2: Segmenting

Issue: Instructions include “You know that each letter has a sound” which may lead to errors by assessor during marking.

Recommendation: Though we agree with SHRP/RTI that “A person who knows how to spell a word is able to access information about a word from two routes, orthographically and phonetically”, it is

unlikely that P1 and P2 students will use an orthographical (spelling) path to segmenting, while assessors may be influenced by the orthography since they are reading the word and may thus make errors when marking. After the NORC/STS team first approached SHRP regarding this issue, quick changes were made to the English/Leblango Tangerine tool. The English instructions were corrected to “You know that each letter has a sound and each word is made up of sounds...Listen to the word then tell me all the sounds in the word”. The corrections to the instructions are reasonable since the focus is now on the sounds in the words (so the “phonetic” route to segmentation) while still allowing for the orthographic route. However, the translation in Leblango (or in other local languages) for this instruction was not corrected though other sections of the instructions for this task were corrected, i.e. “If a child provides a letter name rather than the sound, say please tell me the sounds in the words, not the letters”. We recommend a **review of the instructions to eliminate any possible confusion with letter sounds and a review of the translations in all local languages to ensure they are consistent with the newly corrected instructions.**

Issue: As assessors are rushing through instructions, the examples used for student practice prior to the start of the task are often not well done. For subtask 2, as well as subtasks 1 and 3, it is critical that the assessor do the first example to provide a model to the student. Instead, too often, assessors ask students to do the first example so that a model is not provided before the student practices the skill with examples 2 and 3.

Recommendation: Particular attention needs to be paid to the instructions given by the assessor during training with an **emphasis on modeling of the skill for students with the 1st example.**

Issue: Instructions for the 1st example do not require assessors to say the word to be segmented twice prior to the student’s response. It is important that assessors say each word twice so that students have more opportunity to listen to the word.

Recommendation: **Instructions should be reviewed so that all words used as examples are said twice** prior to the student practice.

Issue: In local languages, the examples provided include words with 3 or more syllables. Students have a difficult time understanding the task to be done and the length of the words adds another layer of complexity.

Recommendation: For future EGRA assessments in new local languages, SHRP should review the word selection for the examples and reduce the length of the words to 2 syllables. Words of 3 or more syllables should be moved to items 6 to 10 thus providing students with more opportunity to get a correct answer with 2 syllable words with items 1 to 5. We note that this issue may underestimate what students know but for consistency purposes, the task should remain the same for the current local languages' assessments.

Subtask 3: Non-word Reading

Issue: Subtask 3 is a difficult task for students. The choice of words, particularly in the first line, impacts how well a student may do. Some words in the first line seem to be too long or particularly difficult.

Recommendation: For future EGRA assessments in new local languages, SHRP should review the words in the first line to ensure that long or difficult words are not included. This will provide students with a greater opportunity to get at least 1 word correct and will thus reduce the number of early stops. We note that this issue may underestimate what students know but for consistency purposes, the task should remain the same for the current local languages' assessments.

Subtask 5: Vocabulary (English)

Issue: The instruction “I will say other words and you show me examples of those words around us” seems to confuse students who aren’t quite sure what “around us” refers to. As an example, since “shoes” tend to be under the table or under the chair, students may not always consider that under the table or chair is “around us” since the shoes are not lined up on the table with the other objects (pencil, paper, book, rubber) used in this activity.

Recommendation: Provide more precise instructions, or limit the area where objects will be available, or continue exercise with students standing up in front of assessor without any furniture between them.

Issue: Culturally, students will not point to a stranger’s shoes. So when asked to show “shoe”, students who aren’t wearing any, will not point to those of the assessor. Furthermore, the use of the singular “shoe” instead of the more common plural form “shoes” is confusing.

Recommendation: Use plural form of “shoes” and not the singular “shoe” which is seldom used. Clarify if assessors should accept as correct when students look under the table at the direction of the assessor’s shoes but do not point.

Issue: Assessors marked as incorrect when students placed the pencil on the right side of the paper but also on the paper. Assessors explained that “to the right of the paper” means on the right side of the paper but on the table (not on the paper).

Recommendation: When a student places the pencil on the right of the paper, whether on the paper or on the table, the answer should be considered correct.

Issue: In part C, the assessor hands the pencil to the student prior to each instruction; the pencil is thereby held by the student in front of his/her body. The assessor takes the pencil away after each prompt and gives it back prior to the next one. When the instruction is “[Put the pencil] in front of you”, students maintain the pencil in front of them. It is unclear however if the lack of reaction is due to the fact that the student understands that the pencil is already in front of him/her (so the student is correct) or due to the student not knowing the answer (so the student is incorrect). Assessors have relied on their own judgment to decide if the student is correct or incorrect.

Recommendation: Assessors should not take the pencil from the student after each prompt to simply hand it back prior to the next prompt. If the pencil remains in the student’s hand the ambiguity will be eliminated as the student will move the pencil from “under the paper” to “in front of you”.

Questionnaire

Issue: Some questions do not have “no response” as an answer choice. Yet students sometimes remain unresponsive. In such cases assessors generally choose an answer, usually no, so they can move forward with the questionnaire.

Recommendation: **Add “no response” as an answer choice to all questions.**

Issue: Students don’t seem to understand the term “library”. They have answered yes to the question regarding taking books home from the library when clearly there isn’t a library at the school.

Recommendation: **Provide a short description of “library” in the question.**

ANNEX 3: MATRIX SUMMARY OF MONTHLY PERFORMANCE FEEDBACK MEMOS

Activity	Constructive feedback provided	Action points (Response from SHRP)	Comments about status of actions suggested.
KAP Assessors training and Fieldwork May to June 2013	<ol style="list-style-type: none"> 1. More time given to the debrief for supervisors held after the pre-test and/or there could be more supervisory training added to the main training. The training focused more on roles and responsibilities of enumerators, less on supervisory responsibilities. 2. Some questions were ambiguous and not clear to the learners and teachers. For example: <ul style="list-style-type: none"> ▪ Q.4 for students: Are you a member of any club where HIV and AIDS are discussed? Does this question refer to membership in clubs at school, outside, or both? ▪ Q.7: Are you a boarding or day learner? Response categories were Yes and No. ▪ Q.8: If a mother has HIV can she pass it on to her baby? Some respondents asked if the mother is enrolled in PMTCT or not, as the risk varies. 3. While data collectors did a good job overall on explaining questions during learner interviews, some did not pause to give learners time to think through their answers. 4. Regarding logistics, some vehicles were old and did not fare well on the roads, and some drivers (e.g., in the pre-test) drove too fast. RTI may wish to consider other companies. 5. Parental concerns and misperceptions (e.g., that their children would be circumcised or tested for HIV against their will) can be better addressed, perhaps with more advance communication or possibly a communication 	<p>Program has taken note and more supervisory training will be added to the main training.</p> <p>The errors are noted and will be fixed, liaising with Evelyn and Stella during tool revision. There is also need to pre-test the revised tool with different groups of learners.</p> <p>Program has taken note of that and will work on it for future trainings.</p>	KAPS follow up data will be collected in October 2015; preparation (assessor training, logistics planning, instrument revision) will take place in the preceding month (September, 2015). These issues will be addressed at this time.

PERFORMANCE & IMPACT EVALUATION (P&IE) SEMI-ANNUAL REPORT

Activity	Constructive feedback provided	Action points (Response from SHRP)	Comments about status of actions suggested.
	that indicates government involvement, e.g., the MOES logo on the consent forms.	<p>Issue noted, there is need to search for different vehicle suppliers. Also design a way to directly pay drivers since the data collectors were also disorganised by unpaid drivers.</p> <p>In the future the program plans to involve SMCs and PTAs to create awareness and mobilise parents to allow their children to participate in KAP survey. In the previous survey, time did not allow this to happen.</p>	
Quarterly report, April 1-June 30, 2013	<p>6. What is your guidance on how best to compare the quarterly reports to the work plan and PMP?</p> <p>7. Review of progress on PMP indicators (May 2013 version)</p> <p>Indicators 2a-2c: when are baseline data expected to be available? e.g.</p> <p>1.1.1. Number of laws, etc. Value is 0. It would be useful to provide an explanation.</p> <p>1.2.2. Number of textbooks. Again value is 0. It would be useful to provide an explanation.</p> <p>1.3.1. Number of teachers. Actual breakouts are very different from expected numbers (more teachers than expected, and fewer CCTs). Is there a reason for that? Or does this not matter? No target is given for 2013, so why are data now being collected? Should there be a target?</p>	<p>The discussion helped SHRP realise disconnect between quarterly reports, PMP, and work plan. They said in the next quarterly report they will seek to provide more clarity on the relationship of the reporting to the work plan and PMP.</p>	<p>The program has developed a “dashboard” which includes all PMP indicators. This is included in the quarterly report and includes a column explaining the status of indicators and reasons for not reaching targets (and planned programmatic modifications).</p> <p>We have redoubled efforts to collect success stories from the field and include these (along with more narrative) in the reports.</p>

PERFORMANCE & IMPACT EVALUATION (P&IE) SEMI-ANNUAL REPORT

Activity	Constructive feedback provided	Action points (Response from SHRP)	Comments about status of actions suggested.
	<p>1.4.1. Number of PTAs or structures supported. Target was 410, actual is zero. Is there a reason?</p> <p>No actual values were provided for a number of indicators (e.g., 1.4.2, 1.5.2, 2.1.1, 2.1.2, 2.2.5 among others). If these are not relevant because, for example, components have been postponed and redesigned, should probably indicate “n/a”.</p>	<p>Program agrees they need to do a better job of telling the story behind the numbers.</p>	<p>All of this information is included in the dashboard.</p>
<p>Rapid Monitoring for HIV and AIDS activities implemented at school.</p> <p>July 2013</p>	<p>8. Some schools did not seem to be clear about the data/recordkeeping requirements during the consultative meetings with school administrators and teachers, it appeared that many R2-related activities have been undertaken at schools but not documented. There appears to be a need to clarify as well as provide reminders to schools on recordkeeping requirements.</p>	<p>As this process is just being rolled out, the program anticipates the on-going need for oversight and training. The program has been exploring options on how to address the issue of data/record-keeping at the school level. One possible solution is to have a journal where teachers and administrators can record interpretation of feedback.</p>	<p>School family initiative, club and guidance and counselling registers are distributed to all program 1651 schools at the beginning of the school year or the end of the last school year. The registers cover a full school year and support our data collection processes. DQ checks have been done on the data from the registers and the results how the quality of the data to be acceptable.</p>
<p>Support supervision for EGR activities, July 2013</p>	<p>9. SHRP Clinical Support Supervision approach emphasizes the importance of the post-observation feedback session where teachers engage in self-assessments and also receive feedback from supervisors and thereby recognise areas of competency and areas for improvement. However, some teams were not organised to present systematic feedback and/or to build from appreciative to constructive feedback. In addition, teachers were busy responding to questions from a number of supervisors and couldn't take notes. Perhaps there is need to have a process before the post-observation sessions where the supervisors can better organize their feedback so it is systematic and ordered from appreciative to constructive; and also a way to document the feedback for teachers to use in future self-reflection and to create a baseline for future support supervision activities.</p>	<p>Program will work out the best way to provide teacher feedback. However, having multiple supervisors observing one teacher was unique in this case because they were simultaneously modelling the support supervision methodology.</p>	<p>The support supervision process has come a long way since it was initiated in 2013. Feedback processes have been streamlined and teachers are equipped with reading journals to take note of the feedback.</p>

Activity	Constructive feedback provided	Action points (Response from SHRP)	Comments about status of actions suggested.
Joint Planning meeting for year 2 work plan, July 2013	10. We would suggest that future work planning meetings include participation of the MoES Planning Unit and Statistics Department which are anticipated consumers of SHRP data. It appears that these officials rarely attend SHRP meetings and activities.	Program agreed with this comment to some extent and said they will try to address it. The program has become involved in the work carried-out by MoES (i.e. sector review, M & E working group meetings) rather than only expecting participation in program activities. The planning unit has been in the Result 2 activities which include incorporating HIV indicators into the national EMIS.	For this work planning cycle, planning with the MoES will take place next week. Various MoES offices were involved in the initial mapping of activities also (NCDC materials production for example). The program is exploring options for carrying forward the work of integrating the HIV indicators into EMIS.
Early Grade Reading Support Supervisions in Kole and Wakiso, August 2013	11. Teachers challenged by the extent to which they are currently involved in lesson plan development. We do not know if this will be resolved when the instructional materials become available. We have seen many errors in lessons, both by native and non-native speakers of the local language. For example, we observed many teachers presenting lessons with spelling errors, in some cases with as many as five out of six vocabulary words in a list spelled incorrectly. While this problem may be substantially resolved when schools receive printed instructional materials, it presents as something that needs to be addressed in the interim. What is the updated estimate of when teaching materials will be distributed to schools? Will the instructional materials include sufficient content for lessons so as to eliminate or at least reduce errors in lesson plans?	Program agreed that teachers are challenged in preparing lessons in their respective local languages. They believe this problem will be reduced when the instructional materials reach schools which include most of the content teachers need to prepare lessons, e.g., vocabulary words, sentence structure, etc. Teachers received a lesson plan template that guides lesson preparation in line with SHRP methodology and national curriculum requirements. Teachers will continuously receive further support from CCTs/School Inspectors trained in providing technical support supervision. Inconsistencies in the newly developed orthographies will be corrected as teachers provide feedback on the instructional materials.	The program (through MoES channels – CCTs) continues to provide in-class support supervision to ensure teachers are better able to deliver the reading curriculum. The materials were available for the majority of training this year and this made a big difference. The hope is that the next round of materials will be ready for the January 2015 training. The content is sufficient for lesson planning and templates of lesson plans are distributed as part of the training.

PERFORMANCE & IMPACT EVALUATION (P&IE) SEMI-ANNUAL REPORT

Activity	Constructive feedback provided	Action points (Response from SHRP)	Comments about status of actions suggested.
Result 2/HIV Data Management, Assessment, and M&E, August 2013	<p>12. The scope of the training was too wide to be covered in 2.5 days. Some major topics were short shifted as a result. The training included a relatively long presentation on FPO job descriptions, orientation to SHRP result 2 activities, and corresponding data collection needs. These consumed a great deal of time and even so did not appear to conclude to the satisfaction of participants. Perhaps only obvious in hindsight, but such large issues/topics should be the focus of separate sessions, and each session organized to focus on fewer and related topics that can be covered in the allocated training timeframe.</p>	<p>Program recognized the comment and will apply this learning to future workshops.</p>	<p>The program strives to ensure that workshop agendas are not overcrowded leaving participants time for genuine learning and reflection.</p> <p>During the teacher training in January the program developed a facilitator's manual that ensured that the participants participated actively but at the same time ensured that the timetable was not too crowded. Feedback from that training was positive. The days were busy but not too busy so that learning did not happen.</p>
Cluster 2 P1 Material Writing Sessions, August 2013.	<p>13. No officials from NCDC/MOES were observed supporting material writing activities. Are they not needed at this stage?</p> <p>14. In addition, one language group had fewer members. It would be good to have equal teams to ease the work of trainers to allocate tasks to the different pairs within a language group. It was observed that members in the group work in pairs. But groups with fewer writers seem to have individuals working on a task alone in order to achieve the same output at end of the day.</p>	<p>NCDC is responsible for recruiting the panel of writers for SHRP materials, and it is understood that NCDC officials will be checking into the sessions but not attending full-time. Program will discuss this further with the Literacy Advisor and also inquire as to whether there were individuals in the different local language groups who worked alone (writers are expected to perform the tasks in groups).</p>	<p>We have tried to get the MoES slots filled and the curriculum specialists to join but unsuccessfully. It is difficult to get ministry officials to recognise their role in the materials development function since they cannot be paid, and the NCDC officials are always busy in other activities. Some panels have fewer people because not everyone that was elected can come, if they are a member of the language board, for instance, or if they are a CCT that their principal will not allow to come away from the college. When positions have fallen vacant due to such constraints we have requested NCDC to mobilise replacements but this too seems one of the most difficult tasks for them to accomplish. For subsequent groups – Cluster 2 P2 and Cluster 3 P1 we agreed to select some individuals who had demonstrated knowledge of the orthographies of their languages and understanding of the methodology, even if</p>

Activity	Constructive feedback provided	Action points (Response from SHRP)	Comments about status of actions suggested.
			they were not writers, and have them join the panels. It has been productive.
Refresher training for TOTs on Cluster I PI materials (teachers guides and primers), September 2013	<p>15. It was not clear whether trainers had presentations they followed during the sessions. Some sessions were not systematically delivered, and many seemed to lack logical order or content. For example, at one station, the session on learner continuous assessment lasted for less than 30 minutes and was taught together with support supervision; however the agenda showed these topics as two separate presentations. At another station, support supervision was held separately but for only 8-10 minutes. Trainers did not appear to be clear on what to present on these two topics.</p> <p>16. Trainers were not prepared to answer questions about the support supervision book. They could not answer the following kinds of questions:</p> <p>How should the book be used? Was one book to be used by all trained teachers? Would SHRP's support supervision book replace the MoES template currently used by head teachers to monitor classroom lessons?</p>	<p>Program has noted this and recognizes the need to do more in the packaging of materials and improve on the sequencing in a systematic way. This has sparked a lot of discussion around the realities of having 5 hours of training not 8.</p> <p>The teachers' guides had the key sessions to be presented. Continuous assessment was in the teachers guides as well. Sessions that were not in the teachers' guides like support supervision books had separate write-ups to guide the presentations. If the trainers did not get it right: we will address it during support supervision. We do realize in the future, we need to outline the training topics more specifically and clearly – writing up more detailed training plans. However, this information is good for us to be more keen on trainers' capability in future trainings.</p>	<p>More detailed training materials (facilitator's guides) have been developed which include objectives and timing for each session. We have made sure that we have more hours to train by limiting logistics and other non-training activities (such as registration). Also, now the teacher guides and learner primers are available in adequate amounts at the training venues. This was not the case in September. This tremendously facilitates training efforts.</p> <p>The use of the SS book has been a topic of discussion in standalone and the on-going teacher trainings – and also demonstrated during SS. It is now in common use.</p>

PERFORMANCE & IMPACT EVALUATION (P&IE) SEMI-ANNUAL REPORT

Activity	Constructive feedback provided	Action points (Response from SHRP)	Comments about status of actions suggested.
	<p>17. Head teacher requests for guidance on key issues to assess when observing teacher during lessons (we understand RTI has a classroom lesson observation tool, were the head teachers supplied with a copy?</p>	<p>In the future RTI will use modular planning and include a one-page list of expectations for each session.</p> <p>Each head teacher was and will be given a copy of the lesson observation tool that SHRP has developed; which has also been reviewed by MoES. We are trying to get away from relying on paper copies since this has not proven to be sustainable in the past. This indicates a need for more focus on this during the TOT.</p> <p>Program is planning to have leadership training in January 2014 for Head Teachers and will provide more guidance on the use of the book/tool during the training.</p>	<p>This has been done.</p> <p>This instrument has been widely shared and utilized during support supervision. The blue books (referenced above) are also now in wide use (rather the relying on the duplication of the form).</p> <p>The books will be discussed again at the leadership training which is being held this month (August, 2014).</p>
<p>Refresher training for teachers on Cluster I PI materials (teacher</p>	<p>18. We observed that during the practice lesson planning session, participants were divided into large groups ranging from 12-15 people which were too large for full</p>	<p>Issue will be addressed in future; trainers will be advised to break teachers in smaller groups.</p>	<p>Teachers are now in groups of no more than 8 teachers.</p>

PERFORMANCE & IMPACT EVALUATION (P&IE) SEMI-ANNUAL REPORT

Activity	Constructive feedback provided	Action points (Response from SHRP)	Comments about status of actions suggested.
guides and primers), September 2013	participation. Some participants dominated the discussion.		
KAP Assessors Training, September 2013	19. The trainees were a combination of new and veteran assessors. The returning assessors were more active and handled more of the questions than the new assessors. In some instances, the veteran assessors seemed to dominate. New assessors may need to have additional training so that they are equally prepared as the returning/veteran assessors and able to contribute fully.	This is a good point that they have also observed in training EGRA assessors, and will be addressed more fully in future training to new assessors.	This will be taken into consideration for the next round of KAPS planning in September, 2015 (see above).
KAP Survey , September 2013	20. The design and guidance on the parental consent form needs additional work. We observed instances where parents did not tick either “agree” or “disagree” regarding consent but still signed the form, rendering the consent form invalid.	Program has taken note of this and will explore more communication channels to reach parents on survey and consent forms. Also consider ways to improve the design and guidance on consent forms in future surveys.	This will be taken into consideration for the next round of KAPS planning in September, 2015 (see above).
EGRA survey data collection, October 2013	21. The assessors consistently made short introductions to prepare learners for interviews. The introduction did not seem to have the desired effect of building confidence of learners to speak up. Some learners remained timid, avoided eye contact with the assessor, and did not seem to listen to the instructions or really read the protocol they were provided. If there are many learners that can in fact read some letters/words but are too shy in front of the assessors to talk, it will affect the data. Overall, the assessors seemed insufficiently skilled in making learners comfortable and attracting and maintaining learner attention. Future trainings should build skills in these areas so that assessors are more adept in conducting interviews with timid learners as this kind of learner will be encountered at all EGRA stages. 22. One of the sub-tasks the assessors involved asking learners to identify objects placed on a table, e.g., a	Building learner rapport is included in the training and emphasized throughout field work. We will look at our training plan to see if there is sufficient emphasis and practice. We believe most of our assessors are very skilled in this area. Of course, when there are outside observers' things always seem more tense on all sides.	This was re-emphasized in the February, 2014 training and the program will continue to do so.

Activity	Constructive feedback provided	Action points (Response from SHRP)	Comments about status of actions suggested.
	<p>pencil, paper, rubber (eraser), etc. While this would seem an easy exercise, it often did not go well. Learners might have been confused by the instructions, many did not speak up, and others said the name but did not point to the object. Some assessors did not place objects on the table. This exercise needs to be reviewed, e.g., to make sure the placing of objects is handled consistently and to address how learners get sufficient information to understand they are supposed to say the word and point to the object.</p>	<p>Learners are never asked to identify objects verbally but to follow English commands (“on the paper”, “behind you”). We of course are, always re-evaluating the use of various tasks and will do so with this one as well based on input from the field but also based on EGRA results. This will also be an emphasis in future training.</p> <p>Program will evaluate the results of this task. P&IE emphasized the main point here is to ensure that the task is consistently carried out the same way by assessors.</p>	<p>Consistency in administration was emphasized in February, 2014 training.</p>
<p>Orientation of C2 Language Board members in Masindi (Runyoro-Rutooro)</p>	<p>23. The two documents distributed during the workshop (scope of work-SOW and workshop program) contained differing objectives, i.e., the SOW stated two objectives (take the LLB members through terms of reference and constitutional framework, and fill vacant positions for LLBs and writing panels), and the workshop program stated three objectives (orient the LLB in its roles and responsibilities; fill vacant positions for LLB and writing panels; and recommend 5 people to work with SIL LEAD to review the orthography). This continues to be confusing and should be easy to remedy.</p> <p>24. During workshop we observed communication issues between the facilitators and the participants. Some participants preferred communicating in the local language but the facilitators selected did not know the local languages. It was quite challenging for facilitators to respond to questions forwarded in the local language or get involved in discussions when participants changed to</p>	<p>We will work to ensure closer alignment between the participants’ documents in the future.</p> <p>Program agrees there should be clear communication between the facilitators and participants when it comes to discussions leading to agreements on critical issues. We understand there is no way participants can carry on productive discussions if the facilitators are not involved.</p>	<p>To allow for adequate engagement in the workshop the numbers of days for the LB meetings was increased to 2 for C3</p> <p>Subsequent LB meetings for had clear objectives consistent with the Program.</p>

Activity	Constructive feedback provided	Action points (Response from SHRP)	Comments about status of actions suggested.
	<p>local language. While it is difficult for RTI to recruit a LLB Consultant who speaks all SHRP languages or to have a Consultant for each region, perhaps the LLB Chairs or some other appropriate person could be enlisted to provide translation during the meetings.</p> <p>25. The workshop had lengthy reading sessions by one facilitator of the different articles in the Constitution, during which participants lost focus and some nodded off. A better way to present the articles of the LLB Constitution is needed to maintain interest and engagement.</p> <p>26. Participants were given two major offsite assignments, localizing the LLB Constitution and developing a work plan for the LLB. Facilitators/organizers assumed these were easy tasks for the group, which eventually turned out to be different. Participants expressed challenges leading themselves through the assignments especially the development of a work plan. In order for RTI to improve on outputs of this activity, there is need to provide participants with written guidance on these tasks and orientation to work plan development as one of the activities on the agenda.</p>	<p>We will note that and ensure that facilitators are equipped with skills to vary presentations. Will see how to borrow interactive techniques used in other SHRP trainings to enrich the delivery of LLB activities.</p> <p>This point is noted. Written guidance will be developed where necessary for these assignments.</p>	<p>We try to have the discussions in English which all the LB members know and understand as a minimum requirement. In cases where there is a switch to LL translations into English will be encouraged to enable the non-local language speakers engage and follow as a matter of procedure.</p> <p>A template for developing a work plan (the major written task mentioned) has been created and distributed to the LB teams.</p>

PERFORMANCE & IMPACT EVALUATION (P&IE) SEMI-ANNUAL REPORT

Activity	Constructive feedback provided	Action points (Response from SHRP)	Comments about status of actions suggested.
Early Grade Literacy Master training on Cluster 1 (P1&P2) and Cluster 2 (P1) materials, December 2013.	27. While a very effective training overall, the registration process was disruptive. Participants were registered and trainers were registered at different times in the training rooms, interrupting sessions.	This is something that will be communicated to the officers supervising the trainings.	Program has noted and attempted to change in more recent trainings. It is often more efficient and manageable to collect attendance information in the classrooms (especially since it has to be collected twice a day) though collecting during the sessions is discouraged.
Early Grade Literacy Training for Trainers on C1 P1 & P2 and C2 P1 SHRP books, December 2013.	<p>28. Lunch meals were always served late, which delayed the afternoon program by an hour or more, and meant that the day's tasks were not completed by the closing time of 5:30 pm. To try to catch up, trainers rushed through Orthography sessions, one of the last items on the daily agenda, and did not have time for participant questions or had to forego Reflection, the last activity on the daily agenda.</p> <p>29. Trainers, especially P2 trainers, need more support in the fourth day of training when using "<i>How to Teach English</i>" materials. There was only one Writer/Expert covering 12 groups. We observed trainers struggling to get clarifications on various instructions and concepts and one trainer misrepresented the SHRP reading model which discourages use of local language in English lessons. We observed a lengthy discussion where participants worried that the English vocabulary word "mama" would confuse learners since it also means mother in local languages; fortunately, the Expert/Writer ultimately visited the class and settled the matter.</p> <p>30. It would be good to cluster administrative announcements and find a way to handle participant</p>	<p>We will try to work more closely with the colleges on this; they are the ones who provide the meals. We might even suggest they hire more staff in order to accommodate the large numbers. In other trainings, we often check in closely with the catering staff and end for lunch only when it is ready.</p> <p>We need to rethink the way we distribute and pair up P2 trainers and assess their skills. In some ways, the trainers have only just been trained themselves.</p>	<p>The colleges have been doing a better job in this regard – feeding upwards of 500 participants at once is not an easy task.</p> <p>Orthography sessions have been taking place in plenary format and facilitators are encouraged to be flexible and work until lunch is ready so as not to lose time.</p> <p>During the refresher training in May, we made effort to address this issue by pairing the more experienced trainers with the less experienced ones. We shall continue pairing the trainers carefully to ensure quality training.</p>

Activity	Constructive feedback provided	Action points (Response from SHRP)	Comments about status of actions suggested.
	<p>registration outside of sessions so as to disrupt sessions less.</p>	<p>This will be communicated to the program training team</p>	<p>The registration process has improved. In most cases participants register before the first sessions, during tea break and lunch time but participant registration of several hundred participants (except for the initial intake) is too difficult to do outside of the classrooms, however facilitators are instructed not to use session time for this.</p>
<p>Early Grade Literacy training for teachers on C1 (P1 & P2) and C2 (P1) SHRP books</p>	<p>31. The training stations were short of training materials (one station had no materials) and participants had to share the few copies available.</p> <p>32. There were inadequate trainers to create manageable groups of participants. At one station, there were more than one hundred participants with only two trainers. In one group, one of the trainers did not speak the local language so the second trainer had to handle most of the load.</p> <p>33. The training content was overly ambitious which meant that some subjects were left out, such as Orthography. Trainers hoped to fit them in elsewhere in the agenda but that wasn't possible.</p>	<p>We are considering developing a training material pack -- or a pack of general instructive materials that will help minimize the risk of collating and printing various sections of our pupil book and teacher guides for the training. This we hope may also assist teachers to get started in the classroom in the event that materials are not available at the beginning of the term. The identification of this pack will eliminate any confusion on the materials that need to go to the training vs. the materials that will be used at the schools.</p> <p>Though this has not been possible to date, it is the hope with Cluster 3 PI materials that they will be available for the training next</p>	<p>It is true, delays in the production of materials hampered more than the ability to get the books into the hands of learners – it negatively impacted the training as materials had not been adequately available.</p> <p>The team in charge of material production has assured us that the teachers' guides and the primers for all clusters will be ready by the time we begin training in January 2015. During the refresher training in May, there were sufficient teachers' guides, pupil books and orthography guides except the Ateso orthography that was being reviewed.</p> <p>We have increased the number of trainers; trainers who do not speak the area local</p>

Activity	Constructive feedback provided	Action points (Response from SHRP)	Comments about status of actions suggested.
		January and then teachers can go right to the classrooms. This will alleviate the need for the development of stop gap materials.	<p>languages are advised to present English lessons.</p> <p>The Situation in May 2014 was much better than 2013. We shall plan manageable content in future trainings. In addition, we shall advise trainers to manage time well because those who do not manage time hinder completion of the planned content.</p>
EGRA C2 Baseline Assessors and Supervisors Training, February 2014	<p>34. Preparation: The training of trainers was very short, primarily consisting of a review of the agenda for first two days of training and watching video of letter sounds. Preparation did not include slides or role-playing. Consequently, there were issues that came up in the training (e.g., transitions between tasks, focus of training modules, etc.) which could have been avoided with adequate training of trainers. Examples:</p> <ul style="list-style-type: none"> a. Assessors were invited to select their own language instrument instead of having a single standard instrument for the first run-through of the tablet version. This caused a lot of confusion because local language versions have a different numbering system than the English version. <p>35. Materials: There were errors or missing instructions on the paper version of the instruments and some inconsistencies with the tablet versions. Errors of this type have the potential to reduce the confidence of assessors in the instruments, so we would recommend allowing adequate time for a careful review before the next round of data collection, taking into consideration the large number of instruments that require review.</p>	<p>One issue is that the English instruments all have instructions that are translated into the various local languages. In fact there is not just one “English version” but 4 and the numbers should all be in agreement between these 4 versions. We will look into this.</p>	<p>The next round of EGRA data will be collected in October, 2014 with training also starting in October. This information will be taken into consideration during the planning for those efforts. As mentioned, there is no one single version – but perhaps we can disable the random function on the tablets so all start with either English or LL.</p> <p>Review of the instruments for the October data collection started in June. Consistency will continue to be the priority.</p>

Activity	Constructive feedback provided	Action points (Response from SHRP)	Comments about status of actions suggested.
	<p>There were no scripts for the demonstrations until the final IRRs. This resulted in some unrealistic and confusing demonstrations, particularly when volunteer assessors carried out the role-play demos.</p> <p>We recommend that the interviewer manual be updated and include a “QxQ” (Question-by-Question) explanation of how each question is to be treated by the interviewer. The manual should be provided to assessors before or at the beginning of training and be considered required reading. Creating and adhering to this manual would reduce conflicting answers provided during training.</p> <p>Training videos were difficult to hear and understand. Both the sound-letter video and the videos of sampling were shown to assessors accompanied by some description. The sampling video was particularly difficult to understand and led to confusion on the part of assessors.</p> <p>36. Activities: Trainers carried out role-plays to demonstrate the assessor-pupil interaction of the main instrument. Except for the final IRRs, these role-plays were not scripted, so they were a bit disorganized and the trainers did not introduce the tasks in advance of the demonstrations clearly. When volunteer assessors were demonstrating, they presented a number of errors. Demonstrating poor techniques before demonstrating good techniques tends to confuse interviewers, so we recommend trainers carry out scripted demos for the first three days and provide scripts to volunteer assessors.</p> <p>The consent/ introduction to the pupil instrument was reviewed quickly on the first day and skipped entirely when the assessors started using the tablets.</p>	<p>Instruments will be reviewed. We aim to do a thorough job of reviewing all instruments and there are in fact differences between paper and Tangerine administration (slashing vs tapping for example). If we are aware of mistakes, we will change them. It should be noted that no assessment were done on paper for the last two data collections. An alternative is to print paper copies directly from Tangerine as is the practice in some other countries. The downside to this is that it uses a considerable amount of paper as the formatting within Tangerine has yet to be optimized.</p> <p>Yes, this is agreed. There was a script (the marked up instrument with previously selected errors) but it is agreed that we need to select volunteers more carefully and then practice the IRRs with more rigor.</p> <p>This seemed to be clear. We believe that seeing and hearing examples is more effective than just speaking. It is not easy to make good videos and we are still working on this.</p>	

Activity	Constructive feedback provided	Action points (Response from SHRP)	Comments about status of actions suggested.
	<p>37. Supervisor training: The field manual made its first appearance at the supervisor training. Supervisors read from a few pages, although trainers quickly noticed that a number of the tasks listed for supervisors were only applicable to paper instruments or were only carried out by DQAs. The entire supervisor training lasted about an hour and there were few opportunities for questions from the new supervisors. We very strongly recommend that a minimum of 2/3 of a day of a well-organized training be dedicated to the supervisor training, as supervisor must guarantee the day to day data quality and logistics of their teams.</p> <p>38. EGRA administration of certain subtasks</p> <p>We noted that certain guidelines given to enumerators about EGRA implementation raised concerns for the impact evaluation. These concerns mainly involve:</p> <ul style="list-style-type: none"> ○ The types of sounds accepted for the letter sound knowledge and segmenting subtasks ○ The types of pronunciation accepted for words in the reading passage <p>We have provided detailed notes regarding the implication of these EGRA implementation guidelines on the impact evaluation in a separate memo that we shared with RTI and USAID on February 21st. NORC and RTI are meeting to discuss the memo on March 18th.2014.</p>	<p>We will work ahead of time to ensure that examples are clearer and that there is enough practice ahead of time.</p> <p>There was a session devoted to this which included a role play. We will look into this but do believe we had ample time to cover everything in our agenda.</p> <p>The supervisor training was not rushed and there was time for questions, it ended at half day and the supervisors remained behind to support packing. We will look into areas that may need more support</p> <p>These issues were discussed on March 18th. SHRP will ensure that the range (narrow) of acceptable letter sounds is well known by the assessors. We are also working on passages to ensure that letters and combinations of letters that are prone to maternal language interferences/transfer will be minimized.</p>	<p>This will be a priority for our DQA (regional supervisor) training which will take place in September.</p> <p>This will be developed for the learner context questions. The paper instrument itself provides ample explanation of the EGRA tasks.</p> <p>This will be a focus of September DQA training.</p> <p>An initial day long DQA/supervisor training will take place in September, prior to the larger training.</p> <p>In August, home office and local literacy advisors will hold a half-day session with facilitators and trainers to ensure that all facilitators/trainers are fluent with this range of letter sounds. Reading passages were developed to minimize the issues around maternal language interference.</p>

ANNEX 4: CHALLENGES TO THE IMPACT EVALUATION, AS PRESENTED IN SEMI-ANNUAL REPORTS PRESENTED IN JUNE 2013, OCTOBER 2013, AND APRIL 2014

1. Result 1: Delays in the implementation of Result 1 continued through October 2013. Although all the trainings have taken place, including refresher TOT and teacher training on Cluster 1 PI materials (teacher guides and primers), these instructional materials were still being distributed to Cluster 1 schools as late as September/October 2013. Our understanding is that materials have not reached all schools at the time of writing this report. In addition, the original plans that included three different treatment arms were modified and treatment was uniform across all schools. Baseline data collection for Cluster 1 was completed successfully in February, and follow-on data collection for Cluster 1 is being fielded among a sub-sample of primary schools. These delays and modifications to the implementation do not pose serious risks to the evaluation at this juncture. We plan to evaluate the impact of the program as it was implemented.

While the implementation changes/delays are not a risk to the evaluation design, an important fact to keep in mind, however, is that we do not expect to see the impacts of the full Result 1 intervention (teacher training and instructional materials) during this first impact analysis, using Oct/Nov 2013 data. However, the Oct/Nov 2013 data will provide us with an opportunity to measure the impact of multiple rounds of teacher training.

2. Result 1: The most recent version of the SHRP PMP indicates that no data will be collected from Cluster 2 in 2016. Going forward with this decision would imply that the impact evaluation for Cluster 2 would only be possible for P1 and P2 but not for P3. Given that Cluster 1 did not receive the full intervention in 2013, Cluster 2 will be the only group that will have a chance to receive three years of full treatment from the beginning of their primary education. The Evaluation Expert already mentioned this omission as a concern to USAID and to the IP as well.
3. Result 1: Data for the second EGRA wave are being collected as we write this report. Initial information from the field indicates low response rates (i.e. low numbers of students are being found) in the schools in the Central Region compared to baseline. We are currently working with the IP to try to address this problem and minimize the risks of having a small sample.
4. Result 2: There are several issues related to sample that have surfaced during the ongoing KAP data collection, which are likely to pose threats to the evaluation of Result 2 activities.

We noted in our first Semi-Annual Report that, it was not possible to include boarding or partial boarding schools -very common among post-primary establishments- in the evaluation sample, given delays in obtaining parental consent for the KAP Survey during the school year. We decided, however, to use the second round of the KAP survey (KAP2) to collect additional baseline data from Cluster 1 boarding and partial boarding post-primary schools by distributing parental consent forms to students before the school break. The idea was to ensure that the baseline survey consisted of a representative sample of post-primary schools, thereby allowing us to generalize the results of the impact evaluation to all such schools in the districts.

We recently learned of several problems that the IP is encountering with the supplemental boarding school component of the second round of KAP surveys. These problems could potentially have serious implications for sample size and the representativeness of the post-primary school sample:

- ▶ The IP faced resistance to data collection activities from some schools, where principals cited concerns that the survey would take away from exam preparation time (national exams in post-primary schools begin in the 2nd week of October) and some head teachers did not distribute consent forms to students at all. These schools could not be interviewed.
- ▶ Some schools closed before the end of the term and consent forms were not distributed on time. These schools could not be interviewed.
- ▶ Other programs related to HIV/AIDS have interacted with some of the schools and, therefore, head teachers decided not to participate in KAP. This is particularly the case of private secondary schools. These schools could not be interviewed.
- ▶ The sample frame that the IP provided NORC for selection of the school sample for the KAP2 contained errors; it included schools that already participated in the first round of KAP. In cases where these schools were randomly selected for the KAP2 sample, they had to be removed from the sample and, where possible, replaced.

NORC has requested from the IP a list of all schools in the KAP2 sample with disposition comments for each of the schools. After evaluating the situation we will have a clearer impression of the effect that these problems can have on the evaluation. At a minimum, we expect a reduction in sample size.

5. Result 2: As mentioned above, SHRP decided not to include post primary establishments in new treatment districts (Cluster 2 and after). Therefore, we will only be able to assess the impact of the Result 2 intervention on post-primary educational facilities for Cluster 1 schools.
6. Result 2: Based on the most recent PMP, we note that the Result 2 intervention will no longer be conducted in Cluster 3 districts and schools. As a result, NORC will focus its evaluation of Result 2 on Cluster 1 and Cluster 2 schools.
7. Result 1: Given program implementation delays in Year 1, the academic term was delayed for one week in the 11 districts of Cluster 1 where the IP is working in order to build in time to prepare and have teacher guides ready for the second training of teachers. Additional classes to compensate for the one week delay are not currently planned. An equivalent delay did not occur in the control district schools; therefore, the academic year in those schools will be one week longer. We do not anticipate a visible effect, but it is worth mentioning how the reality of the program may affect the evaluation.
8. Result 2: After NORC selected the samples for the impact evaluation of the School Health activity, the focus of the intervention underwent some changes in order to align with PEPFAR priorities. We were informed that the intervention would target large schools (with over 150 students) in high HIV prevalence districts; this brought into question the external validity of the impact evaluation and the ability to include non-intervention districts with similar characteristics to treatment districts in the design. However, these new criteria do not seem to have affected the actual selection of districts and we will proceed with the original evaluation design. However the number of treatment schools increased. The IP went ahead with the selection of schools for treatment and control before NORC could approve the selection. As a consequence no replacements for control schools were selected.

This can result in a smaller sample than needed. The Evaluation Expert discussed this issue with the IP and USAID.

April 2014:

1. Result 1: During the Cluster 2 EGRA training and pilot test, the P&IE team observers noted some issues related to the implementation of three specific EGRA subtasks – Letter Sound Knowledge and Word Segmenting, and Oral Passage Reading - and that could have negative implications for the impact evaluation. Annex 3 describes the issues in great detail and also lays out the implications for the impact evaluation. In short, SHRP was using very stringent requirements for accepting letter sounds as correct; for example, while the EGRA toolkit states that “For consonants that can represent more than one sound (i.e., c, g), either answer is acceptable. For vowels, either the short or long sound is accepted (/i/ as in pin or as in pine),” in the SHRP implementation of EGRA only one sound per vowel was being accepted as correct. As well, local pronunciations of words – e.g. “muzzah” for mother – were being marked as incorrect. This raises the concern that learners who actually know correct letter sounds are assessed as not knowing them, since trainers were instructed during training to mark as wrong any very slight deviation from the “ideal” sound of a letter.

This approach can bias the assessment in favor of treatment schools, where students are being taught one correct letter sound or a specific pronunciation of a word, relative to control schools, where a broader set of letter sounds and pronunciations are being taught. We can take as an example the letter B¹: the sound of letter B is /b/ or /buh/². Both sounds are correct and accepted as building skills towards early reading ability. However, the current application of EGRA in Uganda only accepts a perfect clipped sound /b/ as correct. Marking /buh/ as wrong is likely to punish learners in control schools more than it punishes learners in treatment schools, because teachers in treatment schools are trained to teach /b/ as the only correct sound while teachers in control schools are likely to use either /b/ or /buh/ given that both sounds are considered correct. This approach of “teaching to the test” will bias impact findings in favor of treatment schools. NORC is exploring options for measuring this bias in order to adjust impact measures; towards this end, we briefly discussed some alternatives with USAID, such as measuring the bias by conducting experiments to test more and less restrictive versions of EGRA administration.

2. Result 1: Possible contamination of controls. Because the SHRP team is not planning to expand SHRP implementation to additional districts for Cluster 1, they are planning to implement Result 1 activities in control CCTs in the 11 original districts starting in 2014 in order to meet target numbers of trained teachers. However, they plan to exclude the control schools within the control CCTs which were selected for the EGRA data collection and intervene only in the schools from control CCTs which have not been included in the EGRA data collection. Hence, according to the SHRP M&E Team Lead, no teachers in any grade (P1 through P4) in the EGRA control schools will be trained; nor will instructional materials be distributed to these schools.

¹ Similar problems exist with many other consonants such as D, T, P, K, G, etc.

² RTI International, EGRA Toolkit, March 2009
<https://www.eddataglobal.org/documents/index.cfm?fuseaction=pubDetail&ID=149>

CCTs associated with these control clusters will be strictly instructed not to provide any assistance to these control schools.

Strict exclusion of control schools from treatment is critical for the integrity of the impact evaluation design. While SHRP staff has assured us that no control schools will receive any semblance of the Result 1 interventions, we are nonetheless concerned by the possibility of contamination through CCTs or spillover of materials. Any contamination of the control schools will lead to underestimation of the effects of the SHRP Result 1 interventions. We have made this concern clear to both the IP and USAID, and requested that SHRP put in place adequate safeguards to ensure that the control schools in our sample will not be contaminated.

3. Result 1: Non-systematic replacement of sample schools. During Cluster 2 baseline data collection in Mbale district, the SHRP team opted to exclude control schools that use or were presumed to use Luganda and English instead of Lumasaaba as the medium of instruction. The appropriate procedure to replace these schools (following the replacement rule provided) was not followed. Two of these non-Lumasaaba instruction schools were replaced by schools in which the medium of instruction is Lumasaaba; these replacements were picked from the list of preselected schools designated as replacements. The rest of the non-Lumasaaba instruction schools in the district sample were neither assessed nor replaced. We indicated to the IP and USAID that this approach was neither appropriate to keeping the integrity of a random sample nor conducive to comparing SHRP schools to the average public school in Uganda. First, replacing sample schools with hand-picked replacements creates problems with the sample balance. Second, the aim of the evaluation is to assess reading ability of learners in English and local language. While it is not possible to test them in the local language (Lumasaaba, in this case) in schools that do not teach in Lumasaaba, it would still have been possible to test student's performance in English. As such, NORC's Evaluation Expert urged SHRP staff to conduct the EGRA in English in these schools as soon as we learned of the situation. However, the SHRP team did not comply with this request in a timely manner. Therefore, NORC decided that the impact analysis will need to exclude Mbale district altogether.
4. Result 1: Manafwa district is encountering a serious crisis created by teacher transfers in the region. We learned during field observations that most of the teachers trained by SHRP in January 2014 in this region have been transferred to other schools: four of the treatment schools visited by our local staff did not have a trained PI teacher, because s/he had been transferred. It will be critical to have information about the whereabouts of teachers trained by SHRP, since transfers of trained teachers away from treatment schools will have a severe effect on the impact evaluation. If these teachers end up at control schools, the impacts will be even more skewed. We will work with the IP and through our performance evaluation to try and capture the movement of trained teachers between schools.